On the interpretation and identification of simultaneous-equation models

Joosten, Geert

Published in:
Mathematical modelling

DOI:
10.1016/0270-0255(84)90028-9

Mathematical Modelling, Vol. 5, pp. 139–150, 1984
Printed in the USA.
0270-0255/84/$03.00/0

ON THE INTERPRETATION AND IDENTIFICATION OF SIMULTANEOUS-EQUATION MODELS*

GEERT JOOSTEN

University of Groningen
P.O. Box 800
9700 AV Groningen, the Netherlands

(*Received February* 1982)

Abstract—In this paper attention is paid to the interpretation of simultaneous-equation models. By choosing a specific variance-covariance structure of the disturbances each equation can be given a conditional expectation interpretation. In this case there is no identification problem. The relation to the REID-systems (introduced by Wold) and to the Cowles Commission interpretation is pointed out. The ideas are illustrated by means of three simple models.

1. INTRODUCTION

Characteristic for simultaneous-equation models is that explanatory variables occurring in some equation are explained in other equations of the model. The variables for which this is the case are called endogenous variables. A simultaneous-equation model has as many equations as there are endogenous variables. Besides these variables there are usually so-called predetermined variables in the model. Such variables influence the endogenous variables, but they are not influenced by them. Apart from the definition-equations all equations are stochastic.

This study deals with the question how to interpret these models. As is shown the answer to this question is of extreme importance for the identification and thus for the estimation of the parameters of the model.

We restrict ourselves here to linear simultaneous-equation models with nonstochastic predetermined variables. The stochastic variables possess a multivariate normal distribution. Without loss of generality definition-equations are not taken into account.

Two possible interpretations are considered: the conditional-expectation and the linear-combination interpretation. For three simple models the consequences of these interpretations are considered.

2. HISTORICAL BACKGROUND

Interdependent systems are widely used in applied econometric work today. It is therefore astonishing that so little attention is paid to the interpretation of the different equations of such systems. In his classic article written in 1943, Haavelmo [1] writes that "if one assumes that the economic variables considered satisfy, simultaneously, several stochastic relations, it is usually not a satisfactory method to try to determine each of the equations separately from the data, without regard to the restrictions which the other

* An earlier version of this paper was presented at the Third International Conference on Mathematical Modelling, University of Southern California, July 1981.

equations impose upon the same variables" and that "the stochastic properties ascribed to the variables in one of the equations should, naturally, not contradict those that are implied by the other equations." The first statement is essentially a pleading for the use of so-called full-information estimation methods. The second has to do with the meaning that can be given to the different equations. For model 2, discussed in Section 3, Haavelmo [1] shows that the conjecture $E(y_t \mid x_t) = \alpha x_t$ in general is inconsistent with the assumed model specification. Therefore he concludes that "For prediction purposes the original equations of the system have no practical significance, they play only the role of theoretical tools by which to derive the prediction equations," c.q. the reduced form. This means that the different equations cannot be seen as a description of a "controlled experiment", although in the process of constructing the model one acts like that. In this paper it is shown that under the conditional-expectation interpretation such a discrepancy does not occur.

In a series of papers starting in 1959, Wold ([2, 3]) has given much attention to the causal interpretation of multirelation-models. In recursive systems each relation can be given a direct causal interpretation, or each relation is an eo-ipso predictor. An eo-ipso predictor is defined as follows (see Wold [3]):

If a variable y allows the representation $y = f(x_1, \ldots, x_h) + v$ with $E(y \mid x_1, \ldots, x_h) = f(x_1, \ldots, x_h)$, then $f(x_1, \ldots, x_h)$ is called an eo-ipso predictor of y. Structural relations in a simultaneous-equation system do in general not have this property. In the so-called REID (reformulated interdependent) systems, introduced by Wold, the systematic part of each relation is an eo-ipso predictor for the dependent variable of that equation. The REID-system is obtained from the corresponding ID-system by substitution of the explaining endogenous variables in each equation by their expected value, obtained by the reduced form system. In that case the coefficients of the explaining current endogenous variables can be interpreted as reaction-coefficients, not to the actually observed variables but to their expected values.

Under the conditional-expectation interpretation the systematic part of each equation is an eo-ipso predictor for the dependent variable of that equation.

3. THE MODELS

In this paper the following three models are used to illustrate the consequences of both interpretations of simultaneous-equation models considered here.

Model 1

$$y_t = \alpha x_t + \gamma a_t + u_t \tag{1a}$$

$$x_t = \beta y_t + \delta b_t + w_t \tag{1b}$$

Model 2

$$y_t = \alpha x_t + u_t \tag{2a}$$

$$x_t = \beta y_t + w_t \tag{2b}$$

Model 3

$$y_t = \alpha x_t + u_t \tag{3a}$$

$$x_t = \beta y_t + \delta a_t + \gamma b_t + w_t \tag{3b}$$

The models are subjected to the following specifications:

- a_t and b_t are nonstochastic predetermined variables,
- $U_t = \begin{pmatrix} u_t \\ w_t \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{uu} & \sigma_{uw} \\ \sigma_{wu} & \sigma_{ww} \end{pmatrix}\right]$,
- $EU_tU'_{t'} = 0, t \neq t'$,
- $B = \begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}$ is non-singular,
- α, β, γ, δ, σ_{uu}, σ_{uw} and σ_{ww} are fixed unknown parameters,
- y_t and x_t are the endogenous variables of the model; their joint distribution depends on that of u_t and w_t, and
- $t = 1, \ldots, T$. T is the sample size.

4. ON THE INTERPRETATION OF THE MODELS

The question arises how the models introduced in section 2 should be interpreted.

One might say that each equation in a model describes the relation between one endogenous variable (the left side) and the other variables of the model. The specification of model 1 then indicates that b_t has no "direct" influence on y_t. To say it otherwise, (1a) describes the pattern of change of y given (and expressed in) that of the other variables of the model. In this view it is obvious to add the following properties to model 1:

$$E(y_t \mid x_t; a_t, b_t) = \alpha x_t + \gamma a_t$$
$$E(x_t \mid y_t; a_t, b_t) = \beta y_t + \delta b_t$$

and similar properties to the other models considered (Haavelmo [1]). This interpretation will be called the conditional-expectation interpretation. It coincides with the way models are built up.

We may also look at the models in the following way. Rewriting model 1 gives

$$y_t - \alpha x_t = \gamma a_t + u_t$$
$$x_t - \beta y_t = \delta b_t + w_t.$$

This gives rise to the idea that a certain linear combination of the endogenous variables can be written as the sum of a certain linear combination of the exogenous variables and a random variable. Thus

$$E(y_t - \alpha x_t) = \gamma a_t$$
$$E(x_t - \beta y_t) = \delta b_t.$$

This is essentially the model formulation introduced in econometrics by the Cowles Commission [4]. As is pointed out by Hooper [5] and Chow [6] such models can also be treated in terms of canonical correlation. In this paper the last interpretation will be called the linear-combination interpretation.

In the following sections the consequences of both interpretations are examined.

5. THE CONDITIONAL-EXPECTATION INTERPRETATION

5.1. *Model* 1

As a consequence of the conditional-expectation interpretation we have

$$E(y_t \mid x_t; a_t, b_t) = \alpha x_t + \gamma a_t \tag{4a}$$

$$E(x_t \mid y_t; a_t, b_t) = \beta y_t + \delta b_t. \tag{4b}$$

Implicit to these properties is

$$E(u_t \mid x_t; a_t, b_t) = E(w_t \mid y_t; a_t, b_t) = 0. \tag{5}$$

From (1a–b) and the stochastic specification defined in section 2 follows

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2 \left[\begin{pmatrix} \dfrac{\gamma}{1-\alpha\beta} a_t + \dfrac{\alpha\delta}{1-\alpha\beta} b_t \\ \dfrac{\beta\gamma}{1-\alpha\beta} a_t + \dfrac{\delta}{1-\alpha\beta} b_t \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix} \right] \tag{6a}$$

where

$$\begin{pmatrix} \sigma_{uu} & \sigma_{uw} \\ \sigma_{wu} & \sigma_{ww} \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix} \begin{pmatrix} 1 & -\beta \\ -\alpha & 1 \end{pmatrix} \tag{6b}$$

and

$$E \begin{pmatrix} y_t - Ey_t \\ x_t - Ex_t \end{pmatrix} (y_{t'} - Ey_{t'} \quad x_{t'} - Ex_{t'}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad t \neq t'. \tag{6c}$$

This implies

$$y_t \mid x_t \sim N\Bigg(\frac{\sigma_{yx}}{\sigma_{xx}} x_t + \left(\frac{\gamma}{1-\alpha\beta} - \frac{\sigma_{yx}}{\sigma_{xx}} \frac{\beta\gamma}{1-\alpha\beta}\right) a_t + \left(\frac{\alpha\delta}{1-\alpha\beta} - \frac{\sigma_{yx}}{\sigma_{xx}} \frac{\delta}{1-\alpha\beta}\right) b_t, \ \sigma_{yy} - \frac{\sigma_{xy}^2}{\sigma_{xx}}\Bigg) \tag{7a}$$

$$x_t \mid y_t \sim N\Bigg(\frac{\sigma_{xy}}{\sigma_{yy}} y_t + \left(\frac{\beta\gamma}{1-\alpha\beta} - \frac{\sigma_{xy}}{\sigma_{yy}} \frac{\gamma}{1-\alpha\beta}\right) a_t + \left(\frac{\delta}{1-\alpha\beta} - \frac{\sigma_{xy}}{\sigma_{yy}} \frac{\alpha\delta}{1-\alpha\beta}\right) b_t, \ \sigma_{xx} - \frac{\sigma_{xy}^2}{\sigma_{yy}}\Bigg). \tag{7b}$$

It is easily seen that the properties (4a–b) are met in (7a) and (7b) if

$$\frac{\sigma_{yx}}{\sigma_{xx}} = \alpha \quad \text{and} \quad \frac{\sigma_{xy}}{\sigma_{yy}} = \beta. \tag{8}$$

Then

$$\begin{pmatrix} u_t \\ w_t \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\beta}\sigma_{uw} & \sigma_{uw} \\ \sigma_{uw} & -\frac{1}{\alpha}\sigma_{uw} \end{pmatrix} \right]. \tag{9}$$

Under these restrictions the joint distribution of y_t and x_t has the following form

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2 \left[\begin{pmatrix} \pi_{11}a_t + \pi_{12}b_t \\ \pi_{21}a_t + \pi_{22}b_t \end{pmatrix}, \begin{pmatrix} \frac{\pi_{11}}{\pi_{21}}\sigma_{xy} & \sigma_{xy} \\ \sigma_{xy} & \frac{\pi_{22}}{\pi_{12}}\sigma_{xy} \end{pmatrix} \right]. \tag{10}$$

The interpretation considered here therefore implies that model 1 has only five independent parameters: α, β, γ, δ, and σ_{uw}. Between these and the five parameters, π_{11}, π_{12}, π_{21}, π_{22}, and σ_{xy}, of the joint distribution (10) the relations (11a–e) exist.

$$\alpha = \frac{\pi_{12}}{\pi_{22}} \tag{11a}$$

$$\beta = \frac{\pi_{21}}{\pi_{11}} \tag{11b}$$

$$\gamma = \pi_{11} - \frac{\pi_{12}}{\pi_{22}}\pi_{21} \tag{11c}$$

$$\delta = \pi_{22} - \frac{\pi_{12}}{\pi_{11}}\pi_{21} \tag{11d}$$

$$\sigma_{uw} = -\left(1 - \frac{\pi_{12}}{\pi_{22}}\frac{\pi_{21}}{\pi_{11}}\right)\sigma_{xy}. \tag{11e}$$

Because the parameters of the joint distribution of y_t and x_t uniquely determine the parameters of model 1 under the restrictions (8), we can conclude that this model is identified; see e.g. Malinvaud [7, p. 547]. If one wrongly does not take into account restriction (8) then the joint distribution of y_t and x_t has seven parameters. These parameters do not uniquely define the five parameters of model 1; the parameters of the model then are overidentified.

The joint distribution of u_t and x_t can be determined from that of y_t and x_t by

$$\begin{pmatrix} u_t \\ x_t \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y_t \\ x_t \end{pmatrix} - \begin{pmatrix} \gamma & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a_t \\ b_t \end{pmatrix}. \tag{12}$$

Using (10) and (11) we find

$$\begin{pmatrix} u_t \\ x_t \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 & 0 \\ \pi_{21} & \pi_{22} \end{pmatrix}\begin{pmatrix} a_t \\ b_t \end{pmatrix}, \sigma_{xy} \begin{pmatrix} \frac{1}{\beta} - \alpha & 0 \\ 0 & \frac{1}{\alpha} \end{pmatrix} \right] \tag{13}$$

and

$$u_t \mid x_t \sim N\left(0, \left(\frac{1}{\beta} - \alpha\right)\sigma_{xy}\right). \tag{14}$$

The restrictions (8) therefore imply that u_t and x_t are stochastically independent. Analogously w_t and y_t are independent.

5.2. *Model* 2

According to the conditional-expectation interpretation the following properties are introduced

$$E(y_t \mid x_t) = \alpha x_t \tag{15a}$$

$$E(x_t \mid y_t) = \beta y_t. \tag{15b}$$

From (2a–b) and the stochastic specifications introduced in section 2 we get

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix}\right] \tag{16}$$

subject to (6b) and (6c).

From this specification we derive

$$y_t \mid x_t \sim N\left(\frac{\sigma_{yx}}{\sigma_{xx}} x_t, \sigma_{yy} - \frac{\sigma_{xy}^2}{\sigma_{xx}}\right) \tag{17a}$$

$$x_t \mid y_t \sim N\left(\frac{\sigma_{xy}}{\sigma_{yy}} y_t, \sigma_{xx} - \frac{\sigma_{xy}^2}{\sigma_{yy}}\right). \tag{17b}$$

In this case the properties (15a–b) also imply the restrictions (8).

The joint distribution of y_t and x_t then has the following form

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\beta}\sigma_{xy} & \sigma_{xy} \\ \sigma_{xy} & \frac{1}{\alpha}\sigma_{xy} \end{pmatrix}\right]. \tag{18}$$

u_t and w_t are distributed as given in (9) and, as in the previous case u_t, x_t, w_t, and y_t are stochastically independent.

Between the parameters α, β, and σ_{uw} of model 2 and the parameters σ_{yy}, σ_{xy} and σ_{xx} of the joint distribution of y_t and x_t the following relations exist

$$\alpha = \frac{\sigma_{yx}}{\sigma_{xx}} \tag{19a}$$

$$\beta = \frac{\sigma_{xy}}{\sigma_{yy}} \tag{19b}$$

$$\sigma_{uw} = -\left(1 - \frac{\sigma_{xy}^2}{\sigma_{xx}\sigma_{yy}}\right)\sigma_{xy}. \tag{19c}$$

Therefore model 2 is identified under the conditional-expectation interpretation.

5.3. *Model* 3

The conditional-expectation interpretation now demands

$$E(y_t \mid x_t) = \alpha x_t \tag{20a}$$

$$E(x_t \mid y_t) = \beta y_t + \delta a_t + \gamma b_t . \tag{20b}$$

From section 2 follows

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2 \left[\begin{pmatrix} \dfrac{\alpha\delta}{1-\alpha\beta} a_t + \dfrac{\alpha\gamma}{1-\alpha\beta} b_t \\ \dfrac{\delta}{1-\alpha\beta} a_t + \dfrac{\gamma}{1-\alpha\beta} b_t \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix} \right] \tag{21}$$

subject to (6b) and (6c).

The conditional distributions have the following forms

$$y_t \mid x_t \sim N\Bigg(\frac{\alpha\delta}{1-\alpha\beta} a_t + \frac{\alpha\gamma}{1-\alpha\beta} b_t + \frac{\sigma_{yx}}{\sigma_{xx}} \left(x_t - \frac{\delta}{1-\alpha\beta} a_t - \frac{\gamma}{1-\alpha\beta} b_t \right), \sigma_{yy} - \frac{\sigma_{xy}^2}{\sigma_{xx}} \Bigg) \tag{22a}$$

$$x_t \mid y_t \sim N\Bigg(\frac{\delta}{1-\alpha\beta} a_t + \frac{\gamma}{1-\alpha\beta} b_t + \frac{\sigma_{yx}}{\sigma_{yy}} \left(y_t - \frac{\alpha\delta}{1-\alpha\beta} a_t - \frac{\alpha\gamma}{1-\alpha\beta} b_t \right), \sigma_{xx} - \frac{\sigma_{xy}^2}{\sigma_{yy}} \Bigg) . \tag{22b}$$

Substituting (8) in (22a–b) gives (20a–b). Thus also in this case the restrictions introduced by the conditional-expectation interpretation are met by the restrictions $\dfrac{\sigma_{yx}}{\sigma_{xx}} = \alpha$ and $\dfrac{\sigma_{yx}}{\sigma_{yy}} = \beta$. The joint distribution of y_t and x_t is

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2 \left[\begin{pmatrix} \pi_{11} a_t + \pi_{12} b_t \\ \pi_{22} a_t + \pi_{22} b_t \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{xy} \\ \sigma_{xy} & \dfrac{\pi_{21}}{\pi_{11}} \sigma_{xy} \end{pmatrix} \right] \tag{23}$$

where

$$\frac{\pi_{11}}{\pi_{21}} = \frac{\pi_{12}}{\pi_{22}} .$$

The joint distribution of y_t and x_t has five independent parameters which are related to the five parameters of the model in the following way

$$\alpha = \frac{\pi_{11}}{\pi_{21}} = \frac{\pi_{12}}{\pi_{22}} \tag{24a}$$

$$\beta = \frac{\sigma_{xy}}{\sigma_{yy}} \tag{24b}$$

$$\delta = \pi_{21}\left(1 - \frac{\pi_{11}}{\pi_{21}}\cdot\frac{\sigma_{xy}}{\sigma_{yy}}\right) \tag{24c}$$

$$\gamma = \pi_{22}\left(1 - \frac{\pi_{11}}{\pi_{21}}\cdot\frac{\sigma_{xy}}{\sigma_{yy}}\right) \tag{24d}$$

$$\sigma_{uw} = -\left(1 - \frac{\pi_{11}}{\pi_{21}}\cdot\frac{\sigma_{xy}}{\sigma_{yy}}\right)\sigma_{xy}. \tag{24e}$$

As in the previous cases model 3 is also identified under the conditional-expectation interpretation.

6. THE LINEAR-COMBINATION INTERPRETATION

6.1. *Model* 1

The joint distribution of y_t and x_t is in this case as follows

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2\left[\begin{pmatrix} \pi_{11}a_t + \pi_{12}b_t \\ \pi_{21}a_t + \pi_{22}b_t \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix}\right]. \tag{25}$$

The linear-combination interpretation suggests that certain linear combinations of y_t and x_t can be written as deviations from certain linear combinations of a_t and b_t. Rewriting (1a–b) gives

$$\begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \gamma & 0 \\ 0 & \delta \end{pmatrix}\begin{pmatrix} a_t \\ b_t \end{pmatrix} + \begin{pmatrix} u_t \\ w_t \end{pmatrix}. \tag{26}$$

From the equations (27a–c) the relations between the parameters of (25) and those of model 1 can be derived as

$$E(y_t - \alpha x_t) = (\pi_{11} - \alpha\pi_{21})a_t + (\pi_{12} - \alpha\pi_{22})b_t = \gamma a_t \tag{27a}$$

$$E(x_t - \beta y_t) = (\pi_{21} - \beta\pi_{11})a_t + (\pi_{22} - \beta\pi_{12})b_t = \delta b_t \tag{27b}$$

$$\begin{pmatrix} \sigma_{uu} & \sigma_{uw} \\ \sigma_{wu} & \sigma_{ww} \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}\begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix}\begin{pmatrix} 1 & -\beta \\ -\alpha & 1 \end{pmatrix}. \tag{27c}$$

This gives

$$\begin{aligned}
\pi_{11} - \alpha\pi_{21} &= \gamma & \gamma &= \pi_{11} - \frac{\pi_{12}}{\pi_{22}}\pi_{21} \\
\pi_{12} - \alpha\pi_{22} &= 0 & \alpha &= \frac{\pi_{12}}{\pi_{22}} \\
\pi_{21} - \beta\pi_{11} &= 0 & \beta &= \frac{\pi_{21}}{\pi_{11}} \\
\pi_{22} - \beta\pi_{12} &= \delta & \delta &= \pi_{22} - \frac{\pi_{21}}{\pi_{11}}\pi_{12}.
\end{aligned} \tag{28}$$

Between $\{\sigma_{uu}, \sigma_{uw}, \sigma_{ww}\}$ and $\{\sigma_{yy}, \sigma_{xy}, \sigma_{xx}\}$ exists a one-to-one correspondence by the regular matrix B.

The seven parameters of model 1 are uniquely determined by the seven parameters of the joint distribution of y_t and x_t and therefore this model is identified under the linear-combination interpretation. In contrast with the conditional-expectation interpretation the parameters of (25) are not subject to restrictions.

Implicitly we find

$$E(y_t \mid x_t) = \alpha x_t + \gamma a_t + \frac{\sigma_{yx} - \alpha\sigma_{xx}}{\sigma_{xx}}\left(x_t - \frac{\beta\gamma}{1-\alpha\beta}a_t - \frac{\delta}{1-\alpha\beta}b_t\right) \tag{29a}$$

$$E(x_t \mid y_t) = \beta y_t + \delta b_t + \frac{\sigma_{yx} - \beta\sigma_{yy}}{\sigma_{yy}}\left(y_t - \frac{\gamma}{1-\alpha\beta}a_t - \frac{\alpha\delta}{1-\alpha\beta}b_t\right). \tag{29b}$$

6.2. *Model* 2

Under the linear-combination interpretation the joint distribution of y_t and x_t is

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix}\right]. \tag{30}$$

From (2a–b) we get

$$\begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} u_t \\ w_t \end{pmatrix}. \tag{31}$$

The question is which linear combinations of y_t and x_t can be written as deviations from 0. On account of (30) every linear combination can be written in such a way. Therefore α and β are not identified in this case.

6.3. *Model* 3

In this case the joint distribution of y_t and x_t is of the form (25), subject to the restriction $\pi_{11}\pi_{22} - \pi_{21}\pi_{12} = 0$ (see (23)).

Rewriting (3a–b) results in

$$\begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \delta & \gamma \end{pmatrix}\begin{pmatrix} a_t \\ b_t \end{pmatrix} + \begin{pmatrix} u_t \\ w_t \end{pmatrix}. \tag{32}$$

Using (25), (23) and (32) the following relations between the parameters of model 3 and those of the joint distribution of y_t and x_t can be specified

$$\begin{aligned}
&\pi_{11} - \alpha\pi_{21} = 0 \\
&\pi_{12} - \alpha\pi_{22} = 0 \\
&\pi_{21} - \beta\pi_{11} = \delta \\
&\pi_{22} - \beta\pi_{12} = \gamma \\
&\pi_{11}\pi_{22} - \pi_{21}\pi_{12} = 0 \\
&\begin{pmatrix} \sigma_{uu} & \sigma_{uw} \\ \sigma_{wu} & \sigma_{ww} \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}\begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix}\begin{pmatrix} 1 & -\beta \\ -\alpha & 1 \end{pmatrix}.
\end{aligned} \tag{33}$$

From (33) we conclude that α and σ_{uu} can uniquely be determined from the parameters of the joint distribution of y_t and x_t. The remaining parameters are not identified.

In econometric literature α is said to be overidentified because without the restriction $\pi_{11}\pi_{22} - \pi_{21}\pi_{12} = 0$ the parameter α is inconsistently defined by the parameters of the joint distribution of y_t and x_t. The remaining parameters are called underidentified, because there are not enough restrictions on this joint distribution to determine them uniquely.

7. SOME ASPECTS OF ESTIMATION

In econometric theory much attention is paid to the linear-combination interpretation and the estimation problem in that case. We refer to e.g. Malinvaud [7], Theil [8], Johnston [9], Schönfeld [10] and Chow [6].

In this section the estimation of the parameters of the conditional-expectation interpreted models is briefly considered.

As we have seen in section 4 all the parameters of the models considered are identified. This means that the information contained in the joint distribution of y_t and x_t is sufficient to determine these parameters uniquely from those of the joint distribution. In implicit form the maximum-likelihood estimators of these last parameters will be derived.

Model 1

For model 1 the joint distribution of y_t and x_t is given by (10). For simplicity the following reparameterization is used

$$z_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2\left[\begin{pmatrix} s & 1 \\ 1 & r \end{pmatrix}\begin{pmatrix} \pi_{21} & 0 \\ 0 & \pi_{12} \end{pmatrix} c_t, \sigma_{xy}\begin{pmatrix} s & 1 \\ 1 & r \end{pmatrix}\right] \tag{34a}$$

where

$$c_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix}, \quad s = \frac{\pi_{11}}{\pi_{21}}, \quad r = \frac{\pi_{22}}{\pi_{12}}. \tag{34b}$$

The log-likelihood for z_t, $t = 1, \ldots, T$ is

$$L = -T \ln 2\pi - T \ln \sigma_{xy} - \frac{T}{2}\ln(sr - 1) -$$

$$- \frac{1}{2\sigma_{xy}} \sum_{t=1}^{T} \left\{ z_t' \frac{1}{sr - 1}\begin{pmatrix} r & -1 \\ -1 & s \end{pmatrix} z_t - 2c_t'\begin{pmatrix} \pi_{21} & 0 \\ 0 & \pi_{12} \end{pmatrix} z_t \right.$$

$$\left. + c_t'\begin{pmatrix} \pi_{21} & 0 \\ 0 & \pi_{12} \end{pmatrix}\begin{pmatrix} s & 1 \\ 1 & r \end{pmatrix}\begin{pmatrix} \pi_{21} & 0 \\ 0 & \pi_{12} \end{pmatrix} c_t \right\}. \tag{35}$$

Setting the partial derivatives of L to the parameters π_{12}, π_{21}, s, r and σ_{xy} equal to 0 gives

$$r\pi_{12}M_{bb} + \pi_{21}M_{ab} = M_{bx} \tag{36a}$$

$$s\pi_{21}M_{aa} + \pi_{12}M_{ab} = M_{ay} \tag{36b}$$

$$r^2M_{yy} - 2rM_{xy} + M_{xx} = (sr - 1)r\sigma_{xy} + (sr - 1)^2\pi_{21}^2M_{aa} \tag{36c}$$

$$s^2 M_{xx} - 2sM_{xy} + M_{yy} = (sr - 1)s\sigma_{xy} + (sr - 1)^2\pi_{12}^2 M_{bb} \tag{36d}$$

$$rM_{yy} + sM_{xx} = 2(sr - 1)\sigma_{xy} + 2M_{xy} + \pi_{21}(sr - 1)M_{ay} + \pi_{12}(sr - 1)M_{bx} \tag{36e}$$

where, e.g.

$$M_{ab} = \frac{1}{T}\sum_{t=1}^{T} a_t b_t .$$

The solutions $\hat{\pi}_{12}$, $\hat{\pi}_{21}$, $\hat{s}$, $\hat{r}$ and $\hat{\sigma}_{xy}$ of (36a–e) are the maximum-likelihood estimators looked for. Next the maximum-likelihood estimators of the parameters of model 1 are determined by (34b) and (11a–e).

Model 2.

For this model the joint distribution of y_t and x_t is given by (18). From classical multivariate normal distribution theory it is well-known that M_{yy}, M_{xy} and M_{xx} are the maximum-likelihood estimators of $\frac{1}{\beta}\sigma_{xy}$, σ_{xy} and $\frac{1}{\alpha}\sigma_{xy}$. Therefore the maximum-likelihood estimators for α and β are $\frac{M_{xy}}{M_{xx}}$ and $\frac{M_{xy}}{M_{yy}}$. Applying ordinary least-squares to (2a) and (2b) gives the same estimators for α and β.

Model 3.

The joint distribution of y_t and x_t for this model is given by (23), or reparameterized

$$z_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N_2\left[\begin{pmatrix} 1 & 1 \\ p & p \end{pmatrix}\begin{pmatrix} \pi_{11} & 0 \\ 0 & \pi_{12} \end{pmatrix} c_t,\ \sigma_{xy}\begin{pmatrix} q & 1 \\ 1 & p \end{pmatrix}\right] \tag{37a}$$

where

$$c_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix},\quad p = \frac{\pi_{21}}{\pi_{11}} = \frac{\pi_{22}}{\pi_{12}},\quad q = \frac{\sigma_{yy}}{\sigma_{xy}}. \tag{37b}$$

The log-likelihood is in this case is

$$\begin{aligned} L = {} & -T\ln 2\pi - T\ln\sigma_{xy} - \frac{T}{2}\ln(pq - 1) \\ & - \frac{1}{2\sigma_{xy}(pq - 1)}\sum_{t=1}^{T}\left[z_t'\begin{pmatrix} p & -1 \\ -1 & q \end{pmatrix} z_t - \right. \\ & -2c_t'\begin{pmatrix} \pi_{11} & 0 \\ 0 & \pi_{12} \end{pmatrix}\begin{pmatrix} 1 & p \\ 1 & p \end{pmatrix}\begin{pmatrix} p & -1 \\ -1 & q \end{pmatrix} z_t \\ & \left. + c_t'\begin{pmatrix} \pi_{11} & 0 \\ 0 & \pi_{12} \end{pmatrix}\begin{pmatrix} 1 & p \\ 1 & p \end{pmatrix}\begin{pmatrix} p & -1 \\ -1 & q \end{pmatrix}\begin{pmatrix} 1 & 1 \\ p & p \end{pmatrix}\begin{pmatrix} \pi_{11} & 0 \\ 0 & \pi_{12} \end{pmatrix} c_t \right]. \end{aligned} \tag{38}$$

The solution of the following set of equations gives the maximum-likelihood estimators for π_{11}, π_{12}, p, q, and σ_{xy}

$$p\pi_{11}M_{aa} + p\pi_{12}M_{ab} = M_{ax} \tag{39a}$$

$$p\pi_{11}M_{ab} + p\pi_{12}M_{bb} = M_{ba} \tag{39b}$$

$$p(M_{yy} - 2qM_{xy} + q^2M_{xx}) = pq(pq - 1)\sigma_{xy} + (pq - 1)^2(\pi_{11}M_{ax} + \pi_{12}M_{bx}) \tag{39c}$$

$$p^2M_{yy} - 2pM_{xy} + M_{xx} = p(pq - 1)\sigma_{xy} \tag{39d}$$

$$pM_{yy} - 2M_{xy} + qM_{xx} = 2(pq - 1)\sigma_{xy} + (pq - 1)(\pi_{11}M_{ax} + \pi_{12}M_{bx}). \tag{39e}$$

Using (37b) and (24a–e) we get the maximum-likelihood estimators of α, β, γ, δ, and σ_{uw}.

In section 4.1 is shown that the restrictions (4a–b) (and in the same way (15a–b) and (20a–b)) imply that in the models u_t and x_t and also w_t and y_t are independent (see (13)). Because of this the ordinary least-squares estimators of α, β, γ, and δ in the models, interpreted in the conditional-expectation way, are unbiased. However, this method does not take into account the simultaneous structure of the models. Therefore conflicting estimators for σ_{uw} result.

7. CONCLUSIONS

Using the conditional-expectation interpretation there is no problem of identification. This interpretation, however, induces constraints on the variance-covariance structure of the joint distribution of the endogenous variables. The derivation of the maximum-likelihood estimators of the parameters therefore is really troublesome.

The linear-combination interpretation, used in practically all textbooks, leads in some cases to underidentification. We should be aware of the fact that, e.g. for model 1, formula (29a–b) is implicit to this interpretation.

On the contrary, to this last interpretation the equations of the simultaneous models can be given a simple meaning under the conditional-expectation interpretation. In fact they are eo-ipso predictors. This meaning seems to be familiar to the notions from which models are constructed.

REFERENCES

1. T. Haavelmo, The statistical implications of a system of simultaneous equations, in J. W. Hooper and M. Nerlove, *Selected Readings in Econometrics from Econometrica*, Cambridge, Mass., MIT Press (1970).
2. H. Wold, Ends and means in econometric model building, in U. Grenander, ed., *Probability and Statistics, The Harald Cramér Volume*, Stockholm Almquist and Wicksell, (1959).
3. H. Wold, Forecasting by the chain principle, in H. Wold, ed., *Econometric Model Building: Essays on the Causal Chain Approach*, Amsterdam North Holland Publishing Company (1967).
4. W. C. Hood and T. C. Koopmans, eds., *Studies in Econometric Method*, Cowles Commission for Research in Economics, Monograph 14, New York John Wiley and Sons, Inc. (1953).
5. J. W. Hooper, Simultaneous equations and canonical correlation theory, Econometrica **27**, 245–265 (1959).
6. G. C. Chow, A comparison of alternative estimators for simultaneous equations, *Econometrica* **32**, 532–553 (1964).
7. E. Malinvaud, Statistical Methods of Econometrics, Amsterdam, North Holland Publishing Company (1968).
8. H. Theil, Principles of Econometrics, New York, John Wiley and Sons, Inc., (1971).
9. J. Johnston, *Econometric Methods*, New York, McGraw-Hill Book Company, Inc., (1963).
10. P. Schönfeld, *Methoden der Ökonometrie*, Band II, München, Verlag Franz Vahlen, GmbH, (1971).